CDF/PUB/TOP/PUBLIC/11048
version 1.1

Measurement of R=B ($t \to Wb$) / B ($t \to Wq$) in the $t\bar{t}$ dilepton decay channel at CDF

The CDF Collaboration[1]

[1] *URL http://www-cdf.fnal.gov*

(Dated: October 31, 2013)

Abstract

We report a measurement of the ratio of the branching fractions of R = B($t \to Wb$) /B($t \to Wq$) in the dilepton decay channel using data collected with the CDF II detector at the Tevatron, corresponding to an integrated luminosity of 8.7 fb^{-1}. R is obtained using a maximum likelihood estimator. We measured R=0.87± 0.07 (stat+syst). Assuming the unitarity of the CKM matrix and the existence of three quark generations we exctract $\mid V_{tb} \mid$ = 0.93 ± 0.04 (stat+syst).

INTRODUCTION

In the Standard Model of fundamental interactions the top quark decay rates in down-type quaks (d,s,b) are proportional to $\mid V_{tq} \mid^2$, the Cabibbo-Kobayashi-Maskawa (CMK) matrix element that relates the top and the down-type quark. In the hypothesis of three generations of quarks, the unitarity of the CKM leads to $\mid V_{tb} \mid^2 = 0.99915^{+0.00002}_{-0.00005}$, so the top quark decays primarily and almost exclusively to Wb. If this hypothesis is denied, a fourth quark generation would remove the constrains on $\mid V_{tb} \mid$ and lower values would be possible, giving rise to effects in the top properties (cross section, lifetime), but also in CP violation and B mixing. $\mid V_{tb} \mid$ can be measured directly measuring the cross section of single top production, or it can be extracted from the top decay rate in the $t\bar{t}$ sample. The ratio (R) between the branching fraction of top quark decaying to b quark and the branching fractions of the top quark decaying to any kind of down quark is related to $\mid V_{tb} \mid$:

$$R = \frac{\mathcal{B}(t \to Wb)}{\mathcal{B}(t \to Wq)} = \frac{\mid V_{tb} \mid^2}{\mid V_{td} \mid^2 + \mid V_{ts} \mid^2 + \mid V_{tb} \mid^2} \tag{1}$$

If the previous constrains are assumed, R is expected to be $0.99830^{+0.00006}_{-0.00009}$.

Overview

During both Run I and Run II, CDF performed several measurements of R combining the lepton plus jets ($l + jets$) decay channel with the dilepton one, where both the W bosons decay into a lepton and a neutrino. The most recent combined measurement was performed with a luminosity of 162 pb^{-1} and found a value of R$= 1.12^{+0.21}_{-0.19}$(stat)$^{+0.17}_{-0.13}$(syst) extracting R> 0.61 at 95% of CL [1] . Also the DØ collaboration measured R [2] with 5.4 fb^{-1} in the $l + jets$ and the dilepton channels, obtaining R =0.90±0.04 (stat+syst) with R> 0.79 at 95% of CL. Since the old measurement of R was dominated in the error by the low statitics, recently CDF updated the result in the $l + jets$ channel using the complete dataset, up to a luminosity of 8.7 fb^{-1} [3]. The analysis performed a simultaneous fit over R and the top pair production cross section, finding R $= 0.94 \pm 0.09$ (stat+syst) and $\sigma_{p\bar{p}\to t\bar{t}} = 7.5 \pm 1.0$ (stat+syst). Thus, assuming the CKM Matrix unitarity and the three quark generations, a $\mid V_{tb} \mid = 0.97 \pm 0.05$ is found in agreement with the Standard Model prediction. In order to have a complete information from the CDF data, we decided to perform a new measurement of R in the dilepton sample ($t\bar{t} \to W^+ q W^- \bar{q} \to q\bar{q}\bar{\ell}\ell\bar{\nu}\nu$) using the whole dataset.

This analysis is based on the number of b-jets found in $t\bar{t}$ events using the dilepton sample with at least 2 jets in the final state. The charged leptons could be either electrons or muons. Tau leptons are not included. We use SecVtx algorithm, based on the reconstruction of a secondary vertex in the event, in order to identify a jet coming from b-quark fragmentation (b-tagging). Due to the high purity of the $t\bar{t}$ signal in dilepton events it is possible to perform a kinematic measurement of the $t\bar{t}$ cross section. Our strategy is to use this result to predict on the number of $t\bar{t}$ events. We divide our sample in subsets according to dilepton type, number of jets in the final states and events with zero, one or two tags. The comparison between events and the prediction, given by the sum of the expected $t\bar{t}$ estimate and the background yield, in each subsample is made using a likelihood function. Our measured value for R is the one which maximizes the Likelihood, i.e. gives the best match between our expectation and the observed data.

Signal and background evaluation

In the dilepton final state, the $t\bar{t}$ signature consists of two high p_T opposite charged isolated leptons, 2 energeticcontento jets, and large missing transverse energy $\not\!\!E_T$ due to the presence of two escaping neutrinos. The selection we apply is similar to the one used for the measurement of the $t\bar{t}$ cross section described in Ref.[4]. We select events with offline recostructed isolated opposite signed electrons, with $E_T \geq 20$ GeV or muons $p_T \geq 20$ GeV/c. This dataset is purified from other known Standard Model decays with two leptons in the final state by requiring $\not\!\!E_T \geq 25$ GeV (≥ 50 GeV if any lepton or jets is closer than $20°$ from the $\not\!\!E_T$ direction)and high $\not\!\!E_T$ significance for events with dilepton invariant mass in the Z peak. Moreover we require at least two jets of $E_T \geq 20$. To select very energetic events, as a $t\bar{t}$ event is expected to be, the sum of the reconstructed object E_T is required to be greater than 200 GeV.

In our analysis the sample is composed by the $t\bar{t}$ signal and the following backgrounds: irriducible backgrounds for they have naturally the same final state of our signal, which are dibosons (WW, WZ, ZZ) and DY/Z $\to \tau^-\tau^+$, and backgrounds that pass our selection for detector effects, like DY/Z $\to e^-e^+, DY/Z \to \mu^-\mu^+$, where jets could originate from radiation and large $\not\!\!E_T$ from energy mismeasurements, and QCD production of a W with multiple jets with one jet misidentified as a lepton.

The first step of our background estimate is to calculate the Standard Model processes contributions using the Monte Carlo simulation. Then we use data, and appropriate control samples, to estimate contribution due to fakes.

For Dibosons, $DY/Z \to \tau\tau$ the number of background events is calculated using the producion cross sections , the integrated luminosity for dilepton category and the selection acceptance based on MC prediction, corrected for lepton trigger efficiencies, lepton identification scale factor and a factor accounting for differnent jet multiplicity in data and simulations. The Diboson processes are simulated with PYTHIA. NLO cross sections are used(σ_{WW} 11.34 $\pm$ 0.68, σ_{WZ} 3.47 $\pm$ 0.21, σ_{ZZ} 3.62 $\pm$ 0.22)[5]. For DY/Z$\to \tau\tau$ we use ALPGEN + PYTHIA MC simulation. The production cross section is given by the sum of the different contribution of σ_Z corresponding to each different Z + number of jets process. A K-factor of 1.4 is applied[4].

The systematic uncertainty taken into account on MC estimate is a convolution of statistics uncertainty and systematic uncertainty due to lepton identification, jet energy scale and jet multiplicity scale factor.

Drell Yan is one of the largest background in our pretag sample, as it provides naturally two leptons and jets from radiation. Fake $\not\!\!E_T$ can be produced by an energy mismeasurement. Detector effects are difficult to model, therefore we use a data-driven approach outlined in Ref.[4].

Jets faking a lepton can produce a final state with two leptons in several physics processes, the most important in our case is the associated production of W + jets. In order to estimate this source of background we follow a data-driven procedure, based on the probablity of a jet faking a lepton, outlined in [4].

In Tab.I, we give the final results of number of DIL candidates events versus the background and the Standard Model $t\bar{t}$ signal predictions for the full 8.7 fb^{-1} samples. The quoted uncertainties are the sum of the statistical and sistematic uncertainties (here we are considering as sistematics, the cross section uncertainties, the 6% of uncertainty on the luminosity estimates, the errors on the lepton reconstruction and trigger efficiency and on the number of jet scale factors).

For the b-tagged sample estimate we require at least one jet in the event to be b-tagged by SecVtx algorithm. For Standard Model background processes we rely on Monte Carlo estimates to compute their contribution to tagged sample, we consider events passing the

CDF Run II Preliminary, $\mathcal{L}$=8.7 fb^{-1}

Number of pretag events passing the full selection				
Process	ee	$e\mu$	$\mu\mu$	$\ell\ell$
WW	2.8 $\pm$ 0.3	4.3 $\pm$ 0.5	1.2 $\pm$ 0.2	8.4 $\pm$ 1.0
WZ	1.6 $\pm$ 0.2	0.7 $\pm$ 0.1	0.5 $\pm$ 0.1	2.7 $\pm$ 0.3
ZZ	1.0 $\pm$ 0.1	0.30 $\pm$ 0.05	0.45 $\pm$ 0.08	1.7 $\pm$ 0.2
DY$\rightarrow\tau\tau$	2.8 $\pm$ 0.5	4.1 $\pm$ 0.6	1.3 $\pm$ 0.2	8.0 $\pm$ 1.2
DY$\rightarrow ee$, $\mu\mu$	7.9 $\pm$ 1.1	1.5 $\pm$ 0.6	2.5 $\pm$ 0.8	12 $\pm$ 1
Fakes	5.9 $\pm$ 1.9	10.3 $\pm$ 3.4	5.7 $\pm$ 2.1	21.8 $\pm$ 4.3
Total background	21.9 $\pm$ 2.2	21.2 $\pm$ 3.6	11.6 $\pm$ 2.4	54 $\pm$ 7
$t\bar{t}$ (σ=7.4 pb)	70 $\pm$ 6	116 $\pm$ 10	37 $\pm$ 4	223 $\pm$ 20
Total prediction	92 $\pm$ 7	138 $\pm$ 11	49 $\pm$ 4	278 $\pm$ 21
Observed	92	147	47	286

TABLE I. Event summary for the 8.7 fb^{-1} inclusive DIL sample. It is shown the number of background, SM expectation and data candidate events, divided by lepton flavor.

pretag selection with b-tagged jets in Monte Carlo simulation, then we apply the *SecVtx* tagging scale factor of 0.96 $\pm$ 0.05 to take into account differences in the b-tagging efficiencies between data and MC.

For Drell-Yan we have not Drell-Yan $+c\bar{c}$ events generated below the Z mass peak region, so we use only DY/Z + $b\bar{b}$ events to describe DY/Z + Heavy flavor jets background, rescaling the $b\bar{b}$ events to account also for $c\bar{c}$ events.

Events contributing to the b-tagged signal without a real b-quark ("mistag") are due to cases in which the jet is b-tagged but is not coming from an heavy flavour quark. Every jet in an event is analized and assigned a weight. If the jet is b-tagged the SecVtx Scale factor is applied, otherwise we apply the mistag matrix (this assigns to each taggable jet in the event the probability to be mistagged, as a function of event and jet variables (η, E_T)).

We split our sample by the number of b-tags, so for every event the probability of having the requested number of tags is computed and applied as a weight to the event.

To estimate events due to fake leptons we maintained the data driven approach, and apply the fake rate to events $lepton + fakeable$ passing the whole selection and having at least a b-tagged jet.

As we use the silicon detector to measure secondary vertices originated by b-hadrons, we only use the events with fully operational silicon detector.

The background and the signal are evaluated separately for 0, 1 and 2 tags. Tab.II shows the background for 2 jets -1 tag sample, Tab. III shows the backgrounds for 2 jets-2 tags sample. Background for the 0 tag is obtained by subtracting from the pretag background (Tab.I) the 1 and 2 tag estimates. In the same way the 0 b-tag bin in data is the number of pretagged events minus tagged in each category.

CDF Run II Preliminary, $\mathcal{L}$=8.7 fb^{-1}

Number of events with 1 b-tagged jet passing the full selection				
Process	ee	$e\mu$	$\mu\mu$	$\ell\ell$
Dibosons	0.23 ± 0.06	0.30 ± 0.04	0.13 ± 0.04	0.66 ± 0.10
DY+LF	0.90 ± 0.34	0.23 ± 0.12	0.37 ± 0.24	1.50 ± 0.70
DY+HF	0.30 ± 0.08	0.14 ± 0.02	0.18 ± 0.07	0.63 ± 0.12
Fakes	0.93 ± 0.41	2.8 ± 1.1	1.8 ± 0.9	5.6 ± 1.9
Total background	2.4 ± 0.6	3.5 ± 1.1	2.5 ± 0.9	8.3 ± 2.1
$t\bar{t}$ (σ=7.4 pb)	32 ± 3	52 ± 5	16 ± 2	100 ± 9
Total prediction	34 ± 3	56 ± 5	19 ± 2	110 ± 10
Observed	28	52	16	96

TABLE II. Events summary for the 8.7 fb^{-1} for the single tagged DIL sample. It shows the number of background, SM expectation and data candidates, divided by lepton flavor in events with at least 2 jets, but just one b-tagged.

Since our goal is the measurement of the top quark branching ratio to bottom quark, with respect to the total number of top decays for this analysis is crucial the number of b-jets in the event.

CROSS SECTION

Due to the high purity $t\bar{t}$ signal in dilepton events, it is possible to perform a kinematic measurement of $t\bar{t}$ cross section in the pretag sample.

The measured cross section is calculated as:

CDF Run II Preliminary, $\mathcal{L}$=8.7 fb^{-1}

Number of events with 2 b-tagged jets passing the full selection				
Process	ee	$e\mu$	$\mu\mu$	$\ell\ell$
Dibosons	0.021 ± 0.012	0.0043 ± 0.0011	0.0094 ± 0.0056	0.035 ± 0.014
DY+LF	0.015 ± 0.007	0.0051 ± 0.0029	0.0092 ± 0.0062	0.029 ± 0.015
DY+HF	0.11 ± 0.06	0.040 ± 0.006	0.019 ± 0.006	0.17 ± 0.06
Fakes	0.20 ± 0.12	0.34 ± 0.20	0.47 ± 0.40	1.0 ± 0.5
Total background	0.35± 0.13	0.40 ± 0.20	0.51 ± 0.40	1.25 ± 0.53
$t\bar{t}$ (σ=7.4 pb)	9.4 ± 1.3	15 ± 2	4.9 ± 0.7	29 ± 4
Total prediction	9.8 ± 1.4	15.4 ± 2	5.44 ± 0.95	30.8 ± 4.2
Observed	11	18	5	35

TABLE III. Summary of events for the 8.7 fb^{-1} for the double tagged DIL sample. It shows the number of background, SM expectation and data candidates, divided by lepton flavor in events with at least 2 jets and 2 tags.

$$\sigma_{t\bar{t}} = \frac{N_{obs} - N_{bkg}}{A \cdot L} \tag{2}$$

where the denominator is the weighted sum of the corrected acceptance for each dilepton category A_i multiplied by the relative luminosity L_i, N$_{obs}$ is the number of dilepton candidate events, N$_{bkg}$ is the total number of expected background events. The corrected acceptance A$_i$ for the given i-th dilepton category, containing the leptons $\ell_1\ell_2$, is the product of the acceptance for the $t\bar{t}$ ($A_{\ell_1\ell_2}$) simulated by POWHEG NLO MC [6],with $M_{topt} = 172.5$ GeV, and the correction factor that takes into account the lepton trigger efficiencies and reconstruction scale factors.

Using the numbers in Tab.I, we measure $\sigma_{t\bar{t}}$ in the pretag sample:

$$\sigma_{p\bar{p}\to t\bar{t}} = 7.64 \pm 0.55 \text{ (stat)} \pm 0.46 \text{ (lumi) pb.} \tag{3}$$

The measurement is compatibile, within uncertainties, with the CDF cross section measurement in the dilepton channel[4].

MEASUREMENT OF R

We divide our data in bin of dilepton flavor ($ee, e\mu, \mu\mu$) and in bins of tags (0, 1 and 2 tags) so we have a total of 9 independent subsamples where we can compare data to prediction. Our prediction is the total number of events we expect to have in each bin, i.e. the sum of the events due to the background and the ones due to the $t\bar{t}$ signal.

The number of $t\bar{t}$ signal expected for each bin of tags is a function of the probability for a jet to be tagged. This Probability is R dependent since a b quark generated jet is more likely to be b-tagged. Each jet could be b-tagged or not, depending on its flavor and the b-tagging and mistagging efficiencies, so we estimate the average values of these efficiencies, and then use them as inputs in the final fit. We estimate two different contribution to the mistag probability: one for the jets coming from the top and one for radiation jets. We estimate the mistag efficiency for the jet coming from top, taking the average value of the Mistag Matrix (Mistag Matrix fo June 2009) on jets matched to heavy flavor quarks in the $t\bar{t}$ MC, since our MC is generated with R=1. For the mistags due to radiation jets, we take the mistag matrix average contributions on light flavor jets in the $t\bar{t}$ MC sample.

We take into consideration also the contribution due to events with 3 or more jets, considering the fraction of events with 3 or more jets, found in MC simulation.

In each i-th dilepton flavor bin, the number of $t\bar{t}$ events with j tags we expect is:

$$\lambda^{ij} = P(j) N^{i}_{PRETAG} \tag{4}$$

where $P(j)$ is the probability of having j b-tagged jets in the event and N^{i}_{PRETAG} is the number of events expected for the pretag signal in the i-th dilepton flavor bin.

To measure R we compare our predictions to the observed data, using a likelihood function.

LIKELIHOOD FUNCTION

The likelihood function L used in this analysis is:

$$L = \prod_{i} \wp(\mu^{i}_{exp}\ (R, x_j) | N_{obs}) \prod_{j} G(x_i | \bar{x}_j, \sigma_j). \tag{5}$$

In this equation $\wp(\mu^{i}_{exp}\ (R,x_j)|N_{obs})$ is the Poissonian probability to observe in the i-th bin N_{obs} events, given the expected mean value μ_{exp} which is given by the sum of the total backround yields and the $t\bar{t}$ signal.

The index i runs over the 9 combinations of the three dilepton flavors and three tag bins. The Gaussian functions $G(x_i|\bar{x}_j,\sigma_j)$ are functions of the nuisance parameters x_j, with mean $\bar{x}_j$ and variance σ_j and are used to constraint the sources of systematic uncertainties. For example they describe luminosity, background estimates, selection acceptances and relevant efficiencies. In this way we can correlate uncertainties among different channels, using the same parameter for the common source of uncertainty and allowing them to vary with respect to their central values. R is left as a free parameter and we obtain it by performing a maximum likelihood fit to our predictions given observed data. It is chosen the value of R that best matches our prediction with observed data. The equivalent minimization of the -log(L) is done using MINUIT analysis package.

In Fig. 1, we plot the likelihood logarithm shape about its minimum.

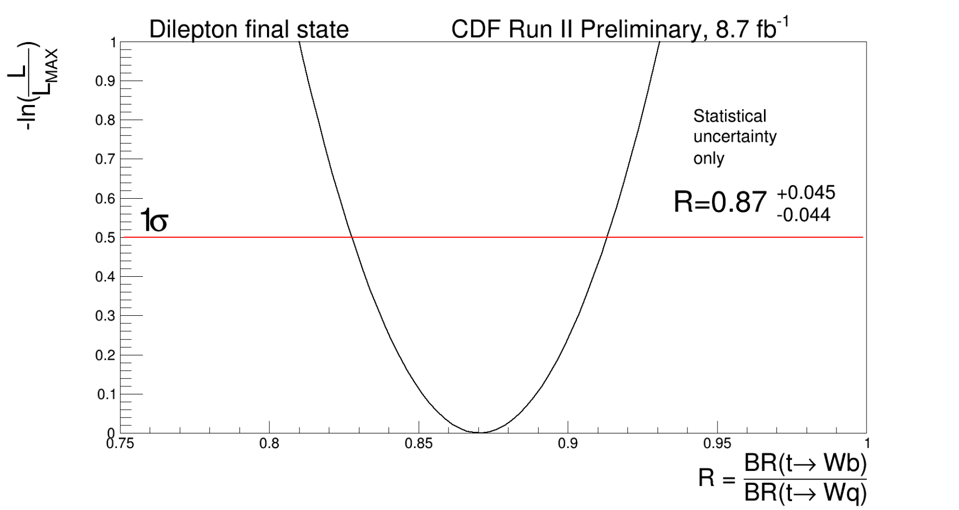

FIG. 1. View of the logarithm likelihood distribution about its minimum.

The overall uncertainty on R is obtained as the value for which $\triangle log(L) = 0.5$. Moreover for the systematics evaluation we considered also the contributions due to the Jet Energy Scale and the initial and final state radiation. The impact of the JES uncertainty is estimated by varying the energy of all jets in the Monte Carlo samples by $\pm\sigma_{JES}$ for both signal and backgrounds. This means a variation on acceptances, cross section and tagging efficiencies.

In order to estimate the ISR/FSR effect, we perform the analysis again using MC samples for the $t\bar{t}$ signal where the ISR/FSR where respectively reduced or enhanced.

The value of R, found with this procedure is :

$$R = 0.871^{+0.074}_{-0.073}\ \textit{(stat+syst)}\ = 0.87 \pm 0.07$$

The uncertainty on R is the combination of statistics and systematics uncertainty. Using our result for R, we obtain a value for the CKM Matrix element: $|V_{tb}| = 0.93 \pm 0.04$.

We show (Fig.2-Fig.5) the number of events found in data and the number of the expected events for different values of R (0.5 (orange), 0.87(blue), 1 (red)), given the background yields plotted in azure, for the different dilepton flavor for 0, 1 and 2 tags.

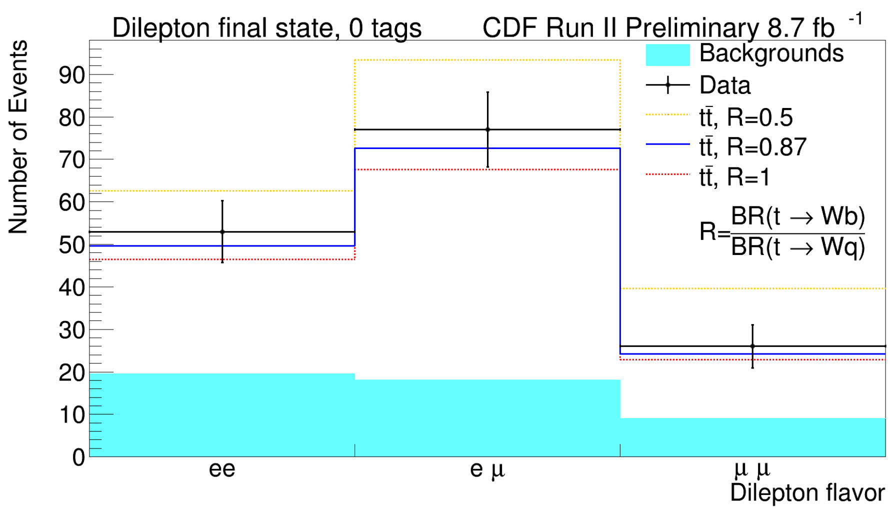

FIG. 2. Number of events found in data and the number of the expected events for several values of R(0.5 (orange), 0.87(blue), 1 (red)), given the background yields plotted in azure. The bins correspond to the dilepton flavors $ee, e\mu, \mu\mu$.

To evaluate the effect of each nuisance parameter on systematics of our result, we performed the fit, varying one by one each nuisance parameter value by a standard deviation from its mean. The most important contributions to R using the procedure are reported in Tab.IV.

CONCLUSIONS

We performed a new measurement of the top quark ratio R, difined as:

$$R = \frac{\mathcal{B}(t \to Wb)}{\mathcal{B}(t \to Wq)} = \frac{\mid V_{tb} \mid^2}{\mid V_{td} \mid^2 + \mid V_{ts} \mid^2 + \mid V_{tb} \mid^2}$$

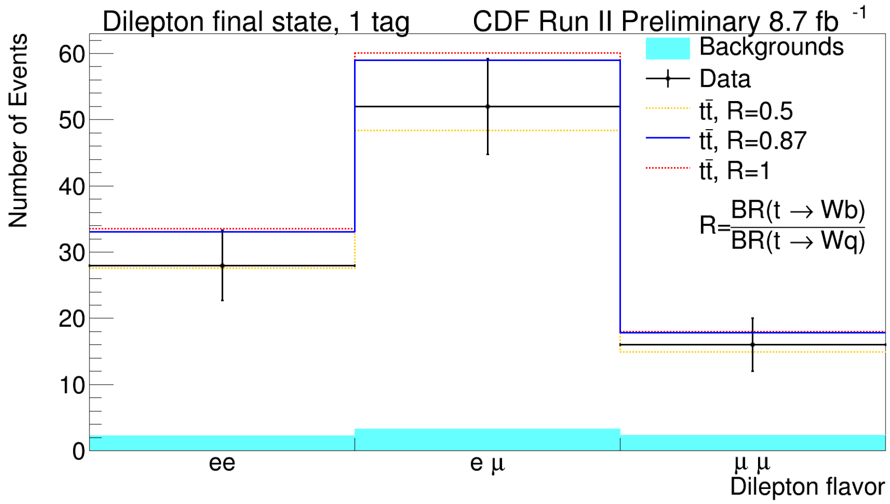

FIG. 3. Number of events found in data and the number of the expected events for several values of R(0.5 (orange), 0.87(blue), 1 (red)), given the background yields plotted in azure. The bins correspond to the dilepton flavors $ee, e\mu, \mu\mu$.

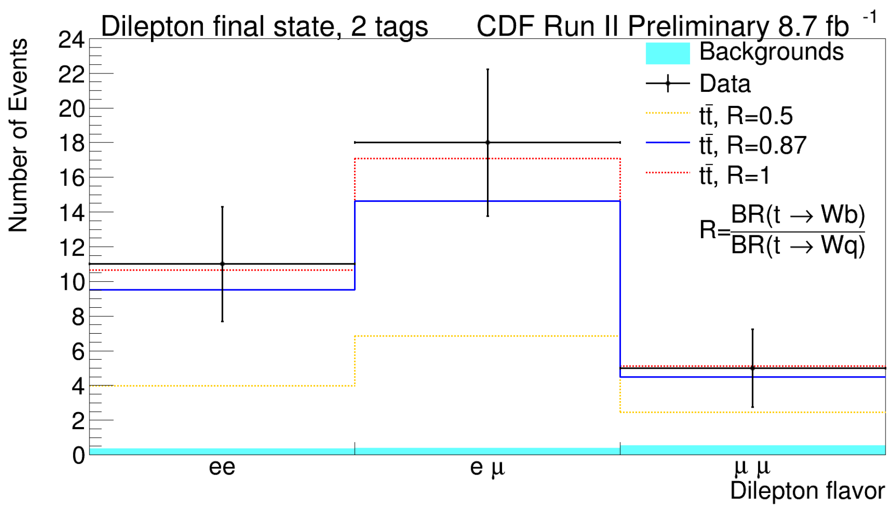

FIG. 4. Number of events found in data and the number of the expected events for several values of R (0.5 (orange), 0.87(blue), 1 (red)), given the background yields plotted in azure. The bins correspond to the dilepton flavors $ee, e\mu, \mu\mu$.

using the dilepton decay channel for the $t\bar{t}$ pair.

We measured R = 0.87 ± 0.07. As a conclusion we provide an estimate of the CKM matrix element that, assuming three generations of quarks and the unitarity of the CKM matrix, is found to be $|V_{tb}|$ = 0.93 ±0.04.

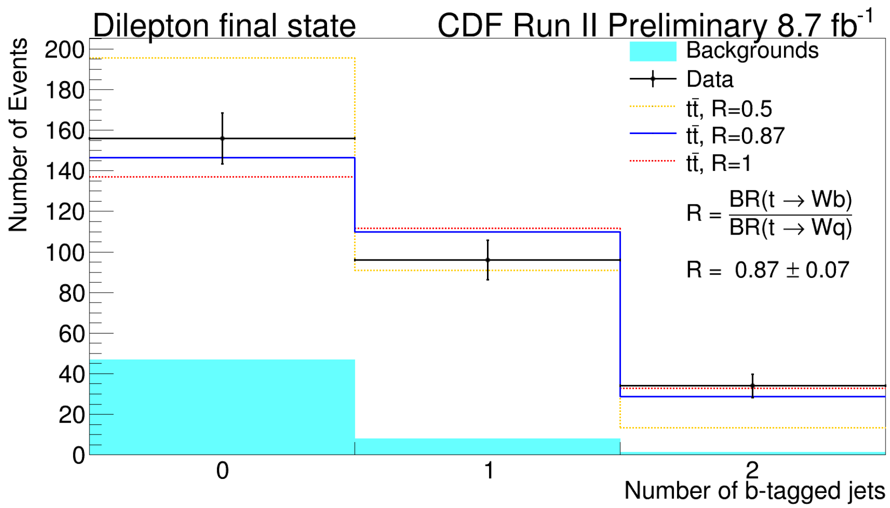

FIG. 5. Number of events found in data and the number of the expected events for several values of R (0.5 (orange), 0.87(blue), 1 (red)), given the background yields plotted in azure. The bins correspond to events with 0,1 and 2 b-tags.

CDF Run II Preliminary, $\mathcal{L}$=8.7^{-1}

Systematic source	Systematic contribution
Correction DATA/MC to b-tagging efficiency	+0.045, -0.040
$\sigma_{t\bar{t}}$	± 0.01
Mistagged jet contribution Luminosity Acceptances , Backgrounds normalization	+0.009, -0.012
Jet Energy Scale	+0.033, -0.025
ISR/FSR	+0.013, -0.025
Total Systematic uncertainty	+0.059, -0.057
Statistical uncertainty	±0.045
Total uncertainty	+0.74, -0.73

TABLE IV. List of the nuisance parameters of the likelihood function that give the largest systematic contribution to our measurement.

[1] S. Acosta *et al.* (CDF Collaboration), Phys. Rev.Lett. **95**, 102002 (1995).

[2] V. M. Abazov *et al.* (D0 Collaboration), Phys. Rev.Lett. **107** (2011).

[3] P. Butti *et al.* (CDF Collaboration), "Measurement of $r = b(t \to wb)/b(t \to wq)$ in the *secvtx* tagged lepton plus jets sample with 8.7 fb^{-1}," (2012), 10887.

[4] C. S. Moon *et al.* (CDF Collaboration), "Measurement of the $t\bar{t}$ production cross-section in events with two leptons and bottom-quark jets using the full CDF data set ," (2013), 10972.

[5] J. M. Campbell and R. K. Elliss, Phys. Rev. D **60**, 113006 (1999).

[6] S. Frixione, P. Nason, and G. Ridolfi, J. High Energy Phys. 09 (2007) 126.